



PCT10

ENTERED

RAW SEQUENCE LISTING

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:41

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

```

3 <110> APPLICANT: Se-Chang KWON
4     Sung-Youb JUNG
5     Sung-Min BAE
6     Gwan-Sun LEE
8 <120> TITLE OF INVENTION: Modified human granulocyte-colony stimulating factor and
process for
9     producing same
11 <130> FILE REFERENCE: PCA00729/HMY
13 <140> CURRENT APPLICATION NUMBER: US/10/031,123B
14 <141> CURRENT FILING DATE: 2002-01-09
16 <160> NUMBER OF SEQ ID NOS: 71
18 <170> SOFTWARE: KOPATIN 1.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 522
22 <212> TYPE: DNA
23 <213> ORGANISM: Homo sapiens
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (1)..(522)
29 <400> SEQUENCE: 1
30 aca ccc ctg ggc cct gcc agc tcc ctg ccc cag agc ttc ctg ctc aag      48
31 Thr Pro Leu Gly Pro Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys
32  1          5          10          15
34 tgc tta gag caa gtg agg aag atc cag ggc gat ggc gca gcg ctc cag      96
35 Cys Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln
36          20          25          30
38 gag aag ctg tgt gcc acc tac aag ctg tgc cac ccc gag gag ctg gtg      144
39 Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val
40          35          40          45
42 ctg ctc gga cac tct ctg ggc atc ccc tgg gct ccc ctg agc tcc tgc      192
43 Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys
44          50          55          60
46 ccc agc cag gcc ctg cag ctg gca ggc tgc ttg agc caa ctc cat agc      240
47 Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser
48 65          70          75          80
50 ggc ctt ttc ctc tac cag ggg ctc ctg cag gcc ctg gaa ggg ata tcc      288
51 Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser
52          85          90          95
54 ccc gag ttg ggt ccc acc ttg gac aca ctg cag ctg gac gtc gcc gac      336
55 Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp
56          100          105          110
58 ttt gcc acc acc atc tgg cag cag atg gaa gaa ctg gga atg gcc cct      384
59 Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro
60          115          120          125

```

RAW SEQUENCE LISTING

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:41

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

```

62 gcc ctg cag ccc acc cag ggt gcc atg ccg gcc ttc gcc tct gct ttc 432
63 Ala Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe
64 130 135 140
66 cag cgc cgg gca gga ggg gtc ctg gtt gct agc cat ctg cag agc ttc 480
67 Gln Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe
68 145 150 155 160
70 ctg gag gtg tcg tac cgc gtt cta cgc cac ctt gcg cag ccc 522
71 Leu Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro
72 165 170
75 <210> SEQ ID NO: 2
76 <211> LENGTH: 174
77 <212> TYPE: PRT
78 <213> ORGANISM: Homo sapiens
80 <400> SEQUENCE: 2
81 Thr Pro Leu Gly Pro Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys
82 1 5 10 15
84 Cys Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln
85 20 25 30
87 Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val
88 35 40 45
90 Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys
91 50 55 60
93 Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser
94 65 70 75 80
96 Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser
97 85 90 95
99 Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp
100 100 105 110
102 Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro
103 115 120 125
105 Ala Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe
106 130 135 140
108 Gln Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe
109 145 150 155 160
111 Leu Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro
112 165 170
115 <210> SEQ ID NO: 3
116 <211> LENGTH: 32
117 <212> TYPE: DNA
118 <213> ORGANISM: Artificial Sequence
120 <220> FEATURE:
121 <223> OTHER INFORMATION: Oligonucleotide primer for the N-terminal of hG-CSF
123 <400> SEQUENCE: 3
124 cgccgccata tgacaccctt gggccctgcc ag 32
127 <210> SEQ ID NO: 4
128 <211> LENGTH: 36
129 <212> TYPE: DNA
130 <213> ORGANISM: Artificial Sequence
132 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:41

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

```

133 <223> OTHER INFORMATION: Oligonucleotide primer for the C-terminal of hG-CSF
135 <400> SEQUENCE: 4
136 accgaattcg gatcctcagg gctgcgcaag gtggcg 36
139 <210> SEQ ID NO: 5
140 <211> LENGTH: 72
141 <212> TYPE: DNA
142 <213> ORGANISM: Artificial Sequence
144 <220> FEATURE:
145 <223> OTHER INFORMATION: Oligonucleotide for preparing E. coli enterotoxin II signal
peptide
147 <400> SEQUENCE: 5
148 tcatgaaaaa gaatatcgca tttcttcttg catctatgtt cgtttttctt attgctacaa 60
150 atgcctacgc gt 72
153 <210> SEQ ID NO: 6
154 <211> LENGTH: 72
155 <212> TYPE: DNA
156 <213> ORGANISM: Artificial Sequence
158 <220> FEATURE:
159 <223> OTHER INFORMATION: Oligonucleotide for preparing E. coli enterotoxin II signal
peptide
161 <400> SEQUENCE: 6
162 acgcgtaggc attttagca atagaaaaaa cgaacataga tgcaagaaga aatgcgatat 60
164 tctttttcat ga 72
167 <210> SEQ ID NO: 7
168 <211> LENGTH: 39
169 <212> TYPE: DNA
170 <213> ORGANISM: Artificial Sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: Oligonucleotide primer coding for the N-terminal of [Ser1]
hG-CSF
175 <400> SEQUENCE: 7
176 acaaatgcct acgcgtctcc cctgggccct gccagctcc 39
179 <210> SEQ ID NO: 8
180 <211> LENGTH: 42
181 <212> TYPE: DNA
182 <213> ORGANISM: Artificial Sequence
184 <220> FEATURE:
185 <223> OTHER INFORMATION: Oligonucleotide primer coding for the C-terminal of [Ser1]
hG-CSF
187 <400> SEQUENCE: 8
188 accgaattcg gatcctcagg gctgcgcaag gtggcgtaga ac 42
191 <210> SEQ ID NO: 9
192 <211> LENGTH: 65
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Oligonucleotide primer coding for E.coli enterotoxin II
Shine-Dalgarno
198 sequence
200 <400> SEQUENCE: 9
201 cggtttcct ctataggttg aggtgtttta tgaaaaagaa tatcgcatTT cttcttgcat 60
203 ctatg 65
206 <210> SEQ ID NO: 10

```

207 <211> LENGTH: 45

RAW SEQUENCE LISTING

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:41

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

208 <212> TYPE: DNA
 209 <213> ORGANISM: Artificial Sequence
 211 <220> FEATURE:
 212 <223> OTHER INFORMATION: Oligonucleotide containing BamHI restriction site
 214 <400> SEQUENCE: 10
 215 accgaattcg gatcctcagg gctgcgcaag gtggcgtaga acgcg 45
 218 <210> SEQ ID NO: 11
 219 <211> LENGTH: 10
 220 <212> TYPE: PRT
 221 <213> ORGANISM: Artificial Sequence
 223 <220> FEATURE:
 224 <223> OTHER INFORMATION: Last five amino acids of E. coli enterotoxin II signal
 peptide plus the
 225 1st to the 5th amino acids of hG-CSF
 227 <400> SEQUENCE: 11
 228 Thr Asn Ala Tyr Ala Thr Pro Leu Gly Pro
 229 1 5 10
 232 <210> SEQ ID NO: 12
 233 <211> LENGTH: 30
 234 <212> TYPE: DNA
 235 <213> ORGANISM: Artificial Sequence
 237 <220> FEATURE:
 238 <223> OTHER INFORMATION: Oligonucleotide for preparing [Thr1]hG-CSF
 240 <400> SEQUENCE: 12
 241 acaaatgcct acgcgacacc cctgggcct 30
 244 <210> SEQ ID NO: 13
 245 <211> LENGTH: 30
 246 <212> TYPE: DNA
 247 <213> ORGANISM: Artificial Sequence
 249 <220> FEATURE:
 250 <223> OTHER INFORMATION: Antisense of SEQ ID NO: 12
 252 <400> SEQUENCE: 13
 253 agggcccagg ggtgtcgcgt aggcatttgt 30
 256 <210> SEQ ID NO: 14
 257 <211> LENGTH: 8
 258 <212> TYPE: PRT
 259 <213> ORGANISM: Artificial Sequence
 261 <220> FEATURE:
 262 <223> OTHER INFORMATION: N-terminal sequence of E. coli enterotoxin II signal peptide
 having
 263 threonine as the 4th amino acid
 265 <400> SEQUENCE: 14
 266 Met Lys Lys Thr Ile Ala Phe Leu
 267 1 5
 270 <210> SEQ ID NO: 15
 271 <211> LENGTH: 33
 272 <212> TYPE: DNA
 273 <213> ORGANISM: Artificial Sequence
 275 <220> FEATURE:
 276 <223> OTHER INFORMATION: Oligonucleotide for substituting the 4th amino acid of E.
 coli
 277 enterotoxin II signal peptide with threonine

RAW SEQUENCE LISTING

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:41

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

```

279 <400> SEQUENCE: 15
280 ggtgttttat gaaaaagaca atcgcatcttc ttc
283 <210> SEQ ID NO: 16
284 <211> LENGTH: 33
285 <212> TYPE: DNA
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Antisense of SEQ ID No: 15
291 <400> SEQUENCE: 16
292 gaagaaatgc gattgtcttt ttcataaaac acc
295 <210> SEQ ID NO: 17
296 <211> LENGTH: 8
297 <212> TYPE: PRT
298 <213> ORGANISM: Artificial Sequence
300 <220> FEATURE:
301 <223> OTHER INFORMATION: C-terminal sequence of E. coli enterotoxin II signal peptide
having
302 glutamine as the 22nd amino acid
304 <400> SEQUENCE: 17
305 Asn Ala Gln Ala Thr Pro Leu Gly
306 1 5
309 <210> SEQ ID NO: 18
310 <211> LENGTH: 26
311 <212> TYPE: DNA
312 <213> ORGANISM: Artificial Sequence
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Oligonucleotide for substituting the 22nd amino acid of E.
coli
316 enterotoxin II signal peptide with glutamine
318 <400> SEQUENCE: 18
319 caaatgccca agcgacaccc ctgggc
322 <210> SEQ ID NO: 19
323 <211> LENGTH: 26
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Antisense of SEQ ID NO: 18
330 <400> SEQUENCE: 19
331 gccaggggt gtcgcttggg catttg
334 <210> SEQ ID NO: 20
335 <211> LENGTH: 24
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <223> OTHER INFORMATION: Oligonucleotide for modifying E. coli enterotoxin II Shine-
Dalgarno
341 sequence
343 <400> SEQUENCE: 20
344 tctagagggt gaggtgtttt atga
347 <210> SEQ ID NO: 21
348 <211> LENGTH: 24
349 <212> TYPE: DNA

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/031,123B

DATE: 11/19/2002
TIME: 14:51:42

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:9; Line(s) 197
Seq#:11; Line(s) 224
Seq#:46; Line(s) 672
Seq#:67; Line(s) 1031
Seq#:69; Line(s) 1070

VERIFICATION SUMMARY

DATE: 11/19/2002

PATENT APPLICATION: US/10/031,123B

TIME: 14:51:42

Input Set : A:\EP.txt

Output Set: N:\CRF4\11192002\J031123B.raw